



OPEN ACCESS

EDITED BY

Luca Steardo Jr,
University Magna Graecia of Catanzaro, Italy

REVIEWED BY

Francesco Monaco,
Azienda Sanitaria Locale Salerno, Italy
Ilaria Pullano,
University of Catanzaro Magna Graecia, Italy

*CORRESPONDENCE

Hemant Ghayvat
✉ hemant.ghayvat@lnu.se

RECEIVED 16 October 2025

REVISED 05 November 2025

ACCEPTED 10 November 2025

PUBLISHED 28 November 2025

CITATION

Tiwari V, Agarwal O, Sharma M, Sahu R,
Babar R, Geddam R, Awais M and Ghayvat H
(2025) Akshar Mitra: a multimodal integrated
framework for early dyslexia detection.
Front. Digit. Health 7:1726307.
doi: 10.3389/fgdth.2025.1726307

COPYRIGHT

© 2025 Tiwari, Agarwal, Sharma, Sahu, Babar,
Geddam, Awais and Ghayvat. This is an open-
access article distributed under the terms of
the [Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Akshar Mitra: a multimodal integrated framework for early dyslexia detection

Vibha Tiwari^{1,2}, Ocean Agarwal^{2,3}, Manya Sharma³, Rashi Sahu³,
Radhika Babar⁴, Rebakah Geddam^{2,5}, Muhammad Awais⁶ and
Hemant Ghayvat^{2*}

¹Center for Artificial Intelligence, Madhav Institute of Technology and Science (Deemed University), Gwalior, India, ²Department of Computer Science and Media Technology, Faculty of Technology, Linnaeus University, Växjö, Sweden, ³Department of Computer Science, Madhav Institute of Technology and Science (Deemed University), Gwalior, India, ⁴Center of Internet of Things, Madhav Institute of Technology and Science (Deemed University), Gwalior, India, ⁵Unitedworld Institute of Technology, Karnavati University, Gujarat, India, ⁶Department of Medical Physics, Memorial Sloan Kettering Cancer Center, New York, NY, United States

Developmental dyslexia is a prevalent neurobiological disorder affecting 10%–15% of children globally, yet it remains largely undiagnosed due to the inaccessibility of conventional assessments in resource-limited settings. Existing screening methods are further constrained by their reliance on unimodal data streams and the need for large, clinically-labeled datasets. This paper presents Akshar Mitra, a Multimodal Integrated Framework (MMF), a novel computational methodology designed for accessible and early dyslexia screening. The framework pioneers the integration of three low-cost, high-yield digital biomarkers derived from eye-tracking, speech, and handwriting analysis. The MMF is implemented through three modules: webcam-based eye-tracking for fixation and saccadic analysis, automated speech assessment for fluency metrics, and optical character recognition for handwriting error detection. Each module extracts 4–6 interpretable features (e.g., fixation regressions, word-error rate, character reversals) that are standardized via a shared data schema. These objective measures are augmented by a concise behavioral questionnaire to generate a holistic risk profile. Beyond screening, the system incorporates support tools, including a dyslexia-friendly reading interface with syllable-level highlighting, to foster user engagement and confidence. By creating a scalable, language-agnostic, and explainable system, this work offers a viable pathway to bridge the global dyslexia diagnostic gap. The MMF provides a transformative tool for proactive screening, facilitating early intervention and improving educational outcomes.

KEYWORDS

dyslexia detection, multimodal framework, early screening, health technology, neurodevelopmental disorders, digital health, cognitive assessment

1 Introduction

Dyslexia (1, 2) is one of the most common learning disorders in children, yet also one of the most misunderstood, particularly in countries such as India. It is characterized by persistent difficulties in reading, spelling, and decoding words, despite adequate intelligence and educational opportunity (3). Globally, dyslexia affects an estimated 5%–15% of the population; in India, recent studies suggest that 10%–15% of school-

aged children i.e., approximately 35–40 million individuals may exhibit dyslexic traits. Unfortunately, more than 80% of dyslexic children in India go undiagnosed and frequently being mislabelled as reckless or underachieving. Such delays in identification can have profound consequences, negatively impacting academic performance, self-esteem, and long-term emotional well-being (4, 25).

The core challenges in early identification stem from limited public awareness, pervasive social stigma, and the high cost of professional assessments (ranging from Rs.5,000 to Rs.30,000), which are primarily offered in urban centers. Most schools, specially in rural and semi-urban regions lack trained personnel and resources to detect learning disabilities at an early stage. Consequently, remedial interventions often occur only after years of struggle, when their efficacy is substantially diminished.

A firm consensus in neuroscience establishes dyslexia as a brain-based learning disorder with a neurobiological origin, completely separate from an individual's general intelligence (5, 6). Neuroimaging research has supplied compelling evidence for structural and functional differences in the brains of individuals with dyslexia compared to typical readers (3). These observable phenomena provide the foundation for the digital biomarkers.

Various technological approaches have also been explored to facilitate the detection and intervention of dyslexia (7). Eye-tracking systems (8), for example, have been used to quantify reading behaviors such as fixation durations and regressions, which differ significantly in dyslexic readers (9). Deep learning-based handwriting analysis has demonstrated promise in detecting dysgraphia and related motor-writing symptoms (10). Similarly, speech-based methods have been employed to assess reading fluency and pronunciation patterns indicative of dyslexia. However, existing solutions are often hardware-dependent, costly, or limited to a single modality, and they seldom account for the socioeconomic diversity encountered in regions like India.

1.1 Contribution and novelty

To address these limitations, **Akshar Mitra**—a unified, AI-powered platform for early dyslexia detection and support—is proposed in this work. Akshar Mitra integrates multiple modalities into a single, accessible system, including:

1. Eye-Tracking Analysis, which captures real-time fixation and saccade metrics via webcam while simultaneously speaking the text through Speech-to-Text Processing, which computes prosodic and fluency features from transcribed reading passages.
2. Handwriting Recognition, using OCR-based extraction of written text followed by symbol-level error analysis
3. Behavioral Quiz, a parental or self-report questionnaire that captures observable dyslexia-related behaviors at home.

Beyond detection, Akshar Mitra offers a dyslexia-friendly *Reading Companion* that incorporates evidence-based practices, large fonts, extra spacing, syllable-level breakdowns, and synchronized audio playback, to reduce cognitive load and enhance reading engagement. This study seeks to make early screening and

intervention more accessible by designing a scalable, affordable, and child-focused solution, with special attention to reaching communities that face barriers due to limited resources or social marginalization. The User Interactive Dashboard of Akshar Mitra is illustrated in **Figure 1**, providing a high-level view of its components and workflow.

To guide the reader through this work, **Section 2** reviews the state of the art in dyslexia detection and highlights key advances in multimodal approaches. **Section 3** then introduces our proposed methodology, which integrates eye-tracking, speech recognition, and OCR-based analysis. **Section 4** reports the experimental validation and results, while **Section 5** concludes with key insights and future research directions.

2 State of the art

The traditional paradigm for the early identification of dyslexia, often reliant on psychometric evaluations and qualitative teacher observations, is undergoing a significant evolution driven by computational intelligence (11, 12). Recent scholarly work highlights a pivotal shift towards data-driven methodologies that offer the potential for more timely and objective screening. Specifically, the application of artificial intelligence (AI) and machine learning (ML) algorithms is enabling the analysis of complex, heterogeneous data to identify subtle neurodevelopmental markers indicative of dyslexia.

To systematically investigate the impact of this technology on dyslexia assessment, this review is structured around three pivotal areas of innovation. The analysis begins by examining modalities that capture fundamental processes impaired in dyslexia: reading and language. We first explore the use of Eye-Tracking for Dyslexia Identification, focusing on how quantitative analysis of saccades, fixations, and regressions provides objective biomarkers of reading difficulties (13, 14). Subsequently, the review investigates the role of Using Speech-to-Text in Dyslexia Diagnosis, delving into how automated analysis of spoken language can reveal underlying phonological deficits.

The scope then broadens to include another critical output skill: writing. This is covered under OCR and AI-Based Handwriting Analysis, which assesses how computational techniques can objectively evaluate graphomotor control and orthographic patterns from written samples. Finally, recognizing that a singular approach may be insufficient, the review culminates with an analysis of Multi-Modal Integration. This section critically evaluates advanced frameworks that combine data from two or more of the aforementioned modalities to enhance diagnostic accuracy and create a more holistic understanding of an individual's neurodevelopmental profile.

2.1 Using eye-tracking for dyslexia identification

The application of eye tracking technology to examine the visual attention and behavior of children offers a timely, objective, and non-invasive way to study developmental



FIGURE 1 Dashboard of Akshar Mitra illustrating its modules.

disorders such as dyslexia. This technology precisely monitors the eyeballs by capturing and monitoring the eye movements through sophisticated cameras. Key eye tracking metrics include fixation counts, processing speed, saccades, scan path, dwell time and gaze duration (15). These measures provide important insights that can effectively differentiate between typical and dyslexic reading behaviors. Coupling machine learning with eye tracking data, they've demonstrated increased classification accuracy and promising possibility for scalable, early dyslexia screening in classrooms. Eye-tracking has also deepened our insight into the psychology of dyslexia by accurately recording eye fixations. Other ML models, including SVMs and CNNs, have also been deployed with great success on eye movement data to predict dyslexics (16, 17). Akshar-Mitra, a proposed solution, incorporates eye-tracking as a fundamental detection feature, likely utilizing it to monitor a child's eye movements during reading tasks to identify patterns indicative of dyslexia (9).

2.2 Using speech-to-text in dyslexia diagnosis

Dyslexia, for instance, is associated with phonological processing deficits—a key factor that can be nicely tested through deep speech analysis. For example, AI-enabled speech recognition might evaluate a child's oral language and detect particular problems with phonological or pronunciation awareness (18, 19). One new diagnostic direction here is to

employ spectrogram analysis with CNNs. In this method, children's raw audio waveforms are converted into spectrogram images which serve as input for CNN architectures. These models have demonstrated superior accuracy in identifying dyslexic individuals compared to traditional diagnostic methods. It shows up in every culture and language, and has similar prevalence in polyglot students as monolingual pupils. However, differentiating "normal" multi-lingual speech delays from an underlying learning disability is challenging and can result in late, over- or under- identification. Bilingual or native language screening is the key to fighting this, assessing such underlying constructs as phonological skills, rapid naming and working memory. While state-of-the-art AI offers the promise of accessing multilingual speech data, this area is also fraught with pitfalls, including language biases and the potential for underrepresented languages to be misrepresented in training data. Akshar-Mitra addresses this using speech recognition, specifically leveraging the OpenAI-Whisper API, to analyze spoken language and identify phonological problems. The system's upcoming updates will also emphasize multilingual and local language support, which is critical for the intended user base.

2.3 OCR and AI-based handwriting analysis

Handwriting analysis using AI has emerged as a viable tool for the early detection of dyslexia and dysgraphia (20). The technology can identify several indicators, including misspelling,

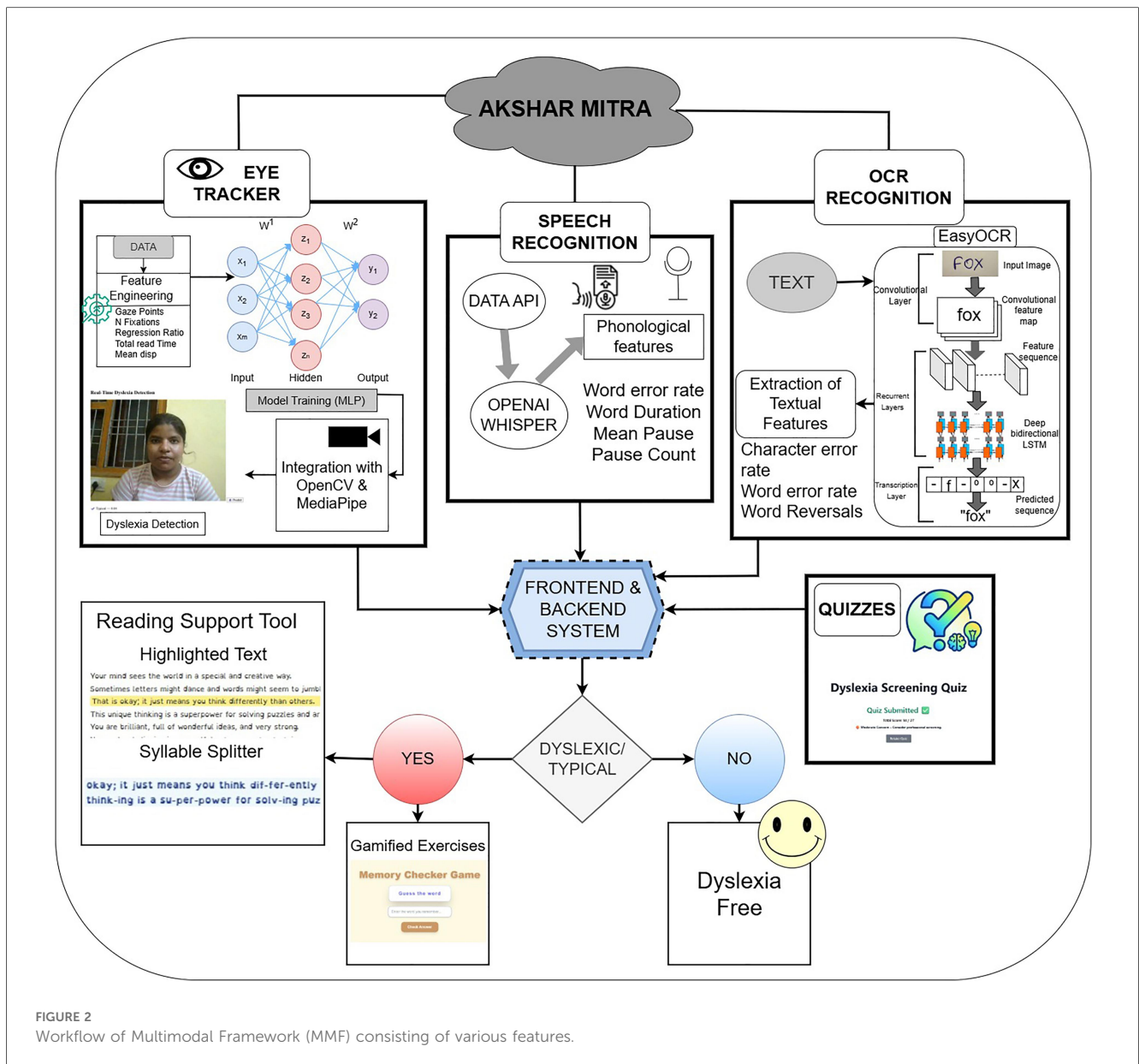


FIGURE 2 Workflow of Multimodal Framework (MMF) consisting of various features.

poor letter formation, and writing disorder. The AI models can identify motor difficulties through writing speed, pressure, and pen movement analysis as well as visual analysis of inconsistencies like irregular spacing and inconsistent letter size, common in people with these conditions. Moreover, AI can read cursive handwriting and digitize it, enable detection of common errors such as misspelling and letter reversal, and even determine underlying cognitive issues by reading grammar and vocabulary. Handwriting has consistently demonstrated unique and subtle features in comparison to the handwriting of dyslexic persons, including letter reversal and motor coordination issues that are amazingly revealing in describing neurological processing. One of the most crucial challenges with present-day AI handwriting analysis models is their transparency, which could make it difficult for teachers and clinicians to understand how specific handwriting features are driving a diagnosis.

However, one of the recurring issues still remains the paucity of diverse handwriting samples of children, which form the crux for effectively training robust AI models, particularly for varying populations.

Akshar-Mitra employs OCR handwriting recognition, with EasyOCR API, to analyze a child’s hand script for specific problems regarding written language processing (21, 26). An integrated multi-modal detection platform bringing together eye-tracking, speech recognition, and handwriting analysis offers a far more complete and perhaps even more accurate diagnostic profile than single-modality methods. This blending encompasses the many-splendored nature of dyslexia symptoms and minimizes dependence on any one, possibly erroneous, data source.

As algorithms for the dyslexia detection become increasingly sophisticated, their applicability in the real world and their effect in regards of fairness are severely constrained by the size,

calibration, and diversity of training data. This underlines a key gap for research: that there is a need for large representative datasets that are culturally and linguistically diverse to harness the maximum potential of AI for the diagnosis of dyslexia internationally, especially in multilingual environments such as India. For AI-based dyslexia screening tools to achieve broad confidence and usage by educators, parents, and clinicians, high accuracy is not enough. These challenges underscore a critical research direction: the development of systems that are not only inclusive but also capable of leveraging multimodal indicators of dyslexia to enhance reliability and generalizability.

3 Materials and methods

This work introduces Akshar Mitra (Framework for Reading, Interpretable Eye-tracking, Dyslexia Screening), a system developed upon the Multimodal Framework (MMF) and is given in Figure 2. The architecture is composed of six interoperable modules, engineered under a unified design to facilitate future multimodal fusion for enhanced screening accuracy.

The system processes three primary data streams in parallel: (1) video for eye-gaze tracking, (2) audio for speech signal analysis, and (3) handwriting samples from digital ink. These core inputs are supplemented by a behavioral assessment module and a reading assistance tool. A dedicated processing pipeline with three parallel encoders performs feature extraction, generating a compact set of 4-6 interpretable features for each data stream. To ensure system-wide compatibility and support real-time inference, all features are standardized into a common schema. This modular architecture is explicitly designed to serve as a foundation for a future unified dyslexia classification model, the details of which are elaborated in subsequent sections.

3.1 Eye-tracking-based fixation analysis

Eye-tracking has been established as a non-invasive and objective method for investigating the neurocognitive processes underlying reading. By analyzing oculomotor metrics during text comprehension, it is possible to identify distinct patterns associated with developmental dyslexia. Empirical evidence consistently demonstrates that individuals with dyslexia exhibit longer fixation durations, a higher frequency of regressive saccades (backward eye movements), and a lower overall reading speed when compared to typical readers. These quantitative metrics serve as robust biomarkers for distinguishing between dyslexic and neurotypical reading behaviors. The integration of these oculomotor features with machine learning algorithms has yielded high-performance classification models, with recent systems achieving detection accuracies of approximately 92% using eye-tracking data alone. This convergence of objective oculomotor analysis and computational modeling provides a powerful paradigm for developing rapid screening tools. Building upon this foundation, this paper introduces a novel

Eye-Tracking-Based Dyslexia Screening Pipeline given as Algorithm 1.

Algorithm 1 Eye-Tracking-Based Dyslexia Screening Pipeline

Require: Folder of fixation CSVs, webcam gaze stream

Ensure: Trained classifier M and real-time dyslexia label \hat{y}

Offline Training:

```

1: function TRAINCLASSIFIER(input_dir)
2:    $F \leftarrow \emptyset$  ▷ initialize empty feature set
3:   for all file in ListFiles(input_dir, "*_fixations.csv") do
4:      $metrics \leftarrow$  EXTRACTFEATURESFROMFILE (file)
5:     if  $metrics \neq$  None then
6:       Append( $F$ ,  $metrics$ )
7:     end if
8:   end for
9:    $D \leftarrow$  DataFrame( $F$ ).DropMissing()
10:  Split  $D$  into  $(X_{train}, y_{train})$  and  $(X_{test}, y_{test})$ 
11:  Normalize features in  $X_{train}, X_{test}$ 
12:  Train MLP classifier  $M$  on  $(X_{train}, y_{train})$ 
13:  Evaluate  $M$  on  $(X_{test}, y_{test})$  using accuracy,  $F_1$ 
14:  return trained model  $M$  and normalizer
15: end function
16: function EXTRACTFEATURESFROMFILE(file_path)
17:  Load CSV into df; if empty, return None
18:  Convert  $duration\_ms$ ,  $disp\_x$ ,  $disp\_y$  to numeric
19:  Compute:
       $n\_fix \leftarrow |df|$ 
       $mean\_fix \leftarrow$  mean( $duration\_ms$ )
       $regressionj\_ratio \leftarrow$  mean( $disp\_x < 0$ )
       $lines\_switches \leftarrow$  df.aoi_line.nunique()
20:   $total\_read\_time \leftarrow$  max( $end\_ms$ ) – min( $start\_ms$ )
21:  return dictionary of metrics
22: end function

```

Online Inference:

```

23: function PREDICTFROMWEBCAM(gaze_points,  $M$ , normalizer)
24:   if  $|gaze\_points| < 5$  then
25:     return error "insufficient data"
26:   end if
27:    $H \leftarrow$  deque(maxlen = 100)
28:   for all pt in gaze_points do
29:     Append( $H$ , pt)
30:   end for
31:    $feats \leftarrow$  ExtractGazeFeatures( $H$ )
32:   if  $feats =$  None then
33:     return error "feature extraction failed"
34:   end if
35:    $feats' \leftarrow$  normalizer.transform( $feats$ )
36:    $\hat{y}, p \leftarrow M.predict\_proba(feats')$ 
37:   return label  $\hat{y}$  with confidence  $p$ 
38: end function

```

The proposed pipeline operationalizes these research principles into a practical and efficient system. It is specifically engineered to utilize a minimalist set of the most salient oculomotor biomarkers, ensuring computational efficiency and robustness for real-time application while maintaining high diagnostic accuracy.

3.1.1 The ETDD70 dataset

This study utilizes the publicly available ETDD70 eye-tracking corpus for model development and validation. The dataset comprises recordings from 70 Czech schoolchildren (aged 9–10), evenly distributed between diagnosed dyslexic and typical reader groups (22). The experimental protocol required participants to read three distinct passages: a syllable matrix, a narrative story, and a pseudo-text. Raw gaze data were processed into fixation and saccade events using the I2MC algorithm, which was selected for its robustness against noise commonly found in pediatric eye-tracking recordings, with a minimum fixation duration threshold of 40 ms (22).

While a comprehensive set of features can be derived, the logistical challenges of high-fidelity eye-tracking make such datasets relatively small, posing a significant risk of model overfitting due to inherent signal variability. Therefore, to ensure model generalizability and interpretability, in this work a stringent feature selection process is adopted, where four trial-level features are identified as having the highest discriminative power between the dyslexic and typical reader cohorts.

3.1.2 Feature selection and justification

A feature selection process was conducted to identify a minimal yet highly discriminative set of metrics for dyslexia classification. From an initial pool of candidates that included fixation variance (*std_fix*) and spatial dispersion (*mean_disp_x*, *std_disp_x*), the final feature set was refined to four key metrics: Fixation Count (*n_fix*), Mean Fixation Duration (*mean_fix*), Regression Ratio (*regression_ratio*), and Total Reading Time (*total_read_time*). The selection was guided by three primary criteria: (i) strong empirical support from prior dyslexia research, (ii) high signal robustness across both controlled and unconstrained webcam-based conditions, and (iii) low redundancy as determined through feature ablation studies. Metrics such as spatial dispersion and line-switch frequency were ultimately discarded due to their high variance and poor generalizability, particularly under the variable lighting conditions inherent to webcam-based eye-tracking.

The selected four-feature set offers significant clinical and computational advantages. Each metric directly corresponds to a well-documented oculomotor characteristic of dyslexia, namely increased fixation frequency, prolonged gaze durations, a higher proportion of regressive saccades, and slower task completion. This minimalist design facilitates stable, real-time inference with minimal computational overhead, ensuring viability in resource-constrained environments like mobile devices or low-resolution video streams while maintaining high classification performance. The resulting feature set thus represents an optimal balance between diagnostic sensitivity, efficiency, and robustness to

signal noise, rendering it highly suitable for scalable deployment in clinical and educational settings.

3.1.3 Machine-learning classification

The classification of the standardized features is performed by a Multilayer Perceptron (MLP). The network architecture comprises a single hidden layer with 64 neurons utilizing a Rectified Linear Unit (ReLU) activation function. The model was optimized using the Adam algorithm with a cross-entropy loss function. For training and evaluation, the dataset was partitioned via stratified sampling into an 80% training set and a 20% held-out validation set to maintain class balance. On the validation data, the classifier achieved an accuracy of approximately 92% and a high F1-score in distinguishing between dyslexic and typical readers. For deployment, the trained model weights and the feature standardization scaler are stored. The inference pipeline first transforms a new feature vector using the saved scaler and then inputs it into the MLP to generate a final dyslexia probability score.

3.1.4 Integrated real-time eye-tracking and speech-based dyslexia screening system

The proposed system integrates a real-time eye-tracking module with a speech analysis component to enable multimodal dyslexia screening during controlled reading sessions. In this setup, a webcam-based gaze tracking application operates concurrently with an audio capture module while the participant reads a displayed sentence aloud.

A continuous video stream is processed on a per-frame basis, where facial landmarks are detected using the MediaPipe, FaceMesh framework to accurately estimate the eye regions. From each processed frame, an approximate gaze point is calculated by averaging predefined eye landmarks corresponding to the pupil and surrounding ocular features. These gaze coordinates are appended to a short-term buffer, typically containing the most recent 100 observations, thereby preserving the temporal evolution of eye movements throughout the reading task.

Once a sufficient number of samples are accumulated, or upon receiving a trigger signal (such as completion of the spoken sentence), the system computes four primary gaze-based features:

1. Number of fixations
2. Mean gaze velocity
3. Count of regressive (backward) saccades
4. Total duration of observation

These features are normalized and passed as input to a pretrained neural classification model, which outputs a probability score indicating the likelihood of dyslexia. The real-time inference result is dynamically overlaid onto the video feed, displaying indicative labels such as “Typical” or “Dyslexic” along with their corresponding confidence values. All feature vectors, predictions, and timestamps are securely logged for subsequent offline analysis. The end-to-end gaze tracking pipeline is optimized for low-latency operation, ensuring interactive performance during live sessions.

Simultaneously, the speech analysis module records and transcribes the participant's spoken output using an Automatic Speech Recognition (ASR) framework. The transcribed text is compared against a predefined reference passage to derive several linguistic and temporal indicators, including:

1. Word Error Rate (WER)
2. Reading speed (words per minute)
3. Frequency and total duration of pauses during reading

Upon completion of the session, the gaze-derived and speech-derived features are jointly analyzed to generate a final dyslexia risk score. This fusion provides a comprehensive assessment by integrating visual attention dynamics with speech fluency and accuracy metrics.

All session data—including raw gaze coordinates, processed features, transcribed speech, and model outputs—are securely stored in an ethically managed institutional database, thereby enhancing the dataset for supervised training and evaluation of future multimodal dyslexia detection models.

This integrated implementation demonstrates the feasibility of real-time, camera-based dyslexia screening that is both portable and non-invasive. By combining gaze movement patterns with speech characteristics, the system provides a reliable framework for early dyslexia risk detection in real-world educational and clinical environments. The detailed operational design and processing pipeline of the speech analysis module are discussed in the following section.

3.2 Speech fluency and timing analysis

The current prototype for speech-based dyslexia screening utilizes an Automatic Speech Recognition (ASR) engine (OpenAI's Whisper) to process recordings of a child reading a fixed text. From the resulting transcription, quantitative features such as Word Error Rate (WER), pause statistics, and speaking rate in Words Per Minute (WPM) are computed against the known ground-truth passage. Dyslexia risk is subsequently assessed using a preliminary, heuristic-based algorithm. While this approach has provided initial validation, its performance is fundamentally constrained by a dependency on these heuristics and the absence of a large, curated dataset for training a robust classification model. The stepwise implementation of this process is formalized in the following [Algorithm 2](#)

Algorithm 2 Speech-Based Dyslexia Feature Extraction and Prediction

Require: Reference transcript, Actual transcript, Word timing list
Ensure: Extracted features, Dyslexia prediction

- 1: Compute **WER** = Word Error Rate between reference and transcript
- 2: Initialize **pauses** list
- 3: **for** $i = 1$ to length of words **do**
- 4: $\text{pause} = \text{words}[i].\text{start} - \text{words}[i - 1].\text{end}$
- 5: **if** $\text{pause} > 0$ **then**

- 6: Append **pause** to **pauses**
 - 7: **end if**
 - 8: **end for**
 - 9: Compute $\text{word_durations} = \text{end} - \text{start}$ for each word
 - 10: Compute $\text{total_time} = \text{last word end} - \text{first word start}$
 - 11: Compute $\text{speech_time} = \text{sum of all word durations}$
 - 12: Calculate features:
 - Mean word duration
 - Mean pause
 - Pause count where $\text{pause} > 0.2 \text{ sec}$
 - Words per minute (WPM) = $\# \text{words} / \text{total_time} \times 60$
 - Articulation rate = $\# \text{words} / \text{speech_time}$
 - 13: Initialize $\text{score} = 0$
 - 14: **if** $\text{WER} \geq 0.15$ **then**
 - 15: $\text{score} \leftarrow \text{score} + 1$
 - 16: **end if**
 - 17: **if** $\text{pause_count} \geq 5$ **then**
 - 18: $\text{score} \leftarrow \text{score} + 1$
 - 19: **end if**
 - 20: **if** $\text{WPM} \leq 80$ **then**
 - 21: $\text{score} \leftarrow \text{score} + 1$
 - 22: **end if**
 - 23: **if** $\text{score} \geq 2$ **then**
 - 24: **return** "Likely Dyslexia"
 - 25: **else**
 - 26: **return** "Typical"
 - 27: **end if**
-

To overcome these limitations, future work is focused on transitioning from the current rule-based system to a data-driven framework. Key objectives include: developing robust methods for precise, forced alignment of the speech signal with the reference text; expanding the feature set to include more informative acoustic and prosodic markers identified in child speech and dyslexia research; and establishing a systematic data collection pipeline, complete with automated report generation, to build a dedicated training corpus.

3.2.1 Child-adaptive ASR and transcription

Children's speech differs acoustically and linguistically from adults'. Thus, using an ASR model adapted to children significantly improves transcription accuracy.

Recent work shows that fine-tuning Whisper on child speech corpora yields lower WER. Jain et al., report that Whisper fine-tuned on 55h of child speech reduces WER on reading tasks.

Fine-tuning Whisper with child-specific corpora (CCHILD) is advisable. Subsequent adaptation can be performed using task-specific labeled data from the target demographic.

3.2.2 Key features for dyslexia screening

The rule-based speech model incorporates a foundational set of features, primarily targeting prosodic and temporal aspects of reading. The feature space is extended to include reading rate, quantified as words per minute (WPM) or syllables per second.

Children at risk of dyslexia often exhibit slower reading speeds accompanied by frequent pauses; this is captured by analyzing aligned timestamps together with pause statistics (e.g., number and duration of pauses longer than 200 ms).

Pronunciation accuracy is modeled through phoneme-level edit distances derived from forced alignment, with the Goodness of Pronunciation (GOP) scoring framework applied to detect mispronunciations. An increased phoneme-error rate is used as an indicator of decoding difficulties.

Rhythmic irregularities are characterized by extracting measures such as pause duration variability and articulation rate. These features provide insight into disrupted prosody, which is a recognized marker of dyslexic speech patterns. While the current focus remains on timing- and pronunciation-based variables, the framework is designed to be extensible. Future extensions may incorporate higher-level prosodic markers such as pitch variation, stress distribution, and phrase boundary detection. Incorporating these variables has demonstrated improvements in both classification performance and interpretability.

3.3 Handwriting OCR and symbolic error evaluation

This module presents an automated pipeline to detect dyslexia-associated writing difficulties using a scanned or photographed image of a handwritten sentence. The algorithm is given as [Algorithm 3](#)

Algorithm 3 OCR-Based Dyslexia Risk Assessment System

Require: Image of handwritten text, Expected sentence
Ensure: Dyslexia risk level, performance metrics

- 1: Initialize EasyOCR reader
- 2: Define reversal pairs: {"b","d"}, {"d","b"}, {"p","q"}, {"q","p"}
- 3: Normalize expected sentence (Unicode-aware lowercase)
- 4: Receive uploaded image and expected sentence from user
- 5: **if** missing input or sentence exceeds 500 words **then**
- 6: **return** Error message
- 7: **end if**
- 8: Save uploaded image temporarily
- 9: Apply OCR to extract text → actual text
- 10: **if** OCR fails to recognize any text **then**
- 11: **return** Error message
- 12: **end if**
- 13: Normalize actual OCR output
- 14: Initialize counters: substitutions, insertions, deletions
- 15: Compare expected vs. actual using SequenceMatcher
- 16: **for all** edit operations (tag, i1, i2, j1, j2) **do**
- 17: **if** tag is "equal" **then**
- 18: Mark characters as correct
- 19: **else**
- 20: Mark characters as incorrect

- 21: Update corresponding counters
- 22: **end if**
- 23: **end for**
- 24: Calculate Character Error Rate (CER)
- 25: Calculate Word Error Rate (WER)
- 26: Count letter reversals using defined pairs
- 27: Compute dyslexia risk score based on:
 - CER and WER
 - Substitutions, Insertions, Deletions
 - Letter reversals
- 28: Classify risk as:
 - **High:** Recommend screening
 - **Moderate:** Monitor further
 - **Low:** No significant indicators
- 29: **return** JSON output with all metrics and risk level

3.3.1 Image processing and OCR

The pipeline uses Optical Character Recognition (OCR) to extract printed text and performs fine-grained comparison against the expected text to detect orthographic inconsistencies. The system is lightweight, deployable in clinical or classroom environments, and capable of real-time analysis and report generation. Users are prompted to write a fixed seed sentence (e.g., "The quick brown fox jumps over the lazy dog") and upload an image of their handwriting via a web interface. The image is processed using *EasyOCR*, a convolutional neural network-based OCR system pre-trained on English scripts. *EasyOCR* returns a linearly ordered character sequence (actual) from the image without needing complex preprocessing, bounding box filtering, or layout correction, making it robust to the variability present in children's handwriting.

3.3.2 Normalization and sequence alignment

Both the reference sentence (expected) and the OCR output are normalized using Unicode normalization (NFKC) and lowercased to reduce orthographic variance. The resulting strings are compared using a dynamic sequence alignment algorithm (*difflib.SequenceMatcher*) that generates character-level edit operations: substitutions, insertions, and deletions. The goal is to identify discrepancies in symbol representation attributed to handwriting errors, OCR artifacts, or spelling problems.

Two primary distance metrics are computed to quantify the deviation from the reference sentence: the Character Error Rate (CER) and the Word Error Rate (WER). The CER is defined as:

$$\text{CER} = \frac{\text{substitutions} + \text{insertions} + \text{deletions}}{\text{number of characters in reference}}$$

This metric captures character-level deviations. The WER, on the other hand, is calculated by tokenizing both the reference and hypothesis texts and computing the word-level edit distance between them. These metrics provide a quantitative assessment

of transcription fidelity and are utilized as core features in the final dyslexia screening decision.

3.3.3 Symbol reversal detection

An additional dyslexia-specific feature is the **letter reversal count**, defined as the number of character pairs that match common visual confusions such as:

$$(b \leftrightarrow d), \quad (p \leftrightarrow q), \quad (q \leftrightarrow p)$$

Such reversals are hallmark indicators of visual-spatial confusion and dysgraphia in early dyslexic children. They are detected via a character-wise scan of the aligned strings and contribute heavily to the screening score.

3.3.4 Rule-based risk scoring

Based on literature-derived thresholds and empirical tuning, a weighted scoring model is implemented. For each sample, the following features are computed: substitution count, insertion count, deletion count, reversal count, character error rate (CER), and word error rate (WER). Each feature contributes one or two points depending on the severity of deviation.

The total score is then mapped to one of three discrete risk classes. A score greater than or equal to 7 (23) indicates high risk, suggesting that immediate and detailed screening is recommended. Scores between 4 and 6 correspond to a moderate risk category, where monitoring for patterns and retesting is advised. A score below 4 reflects a low risk assessment, implying that no significant dyslexia indicators were found.

This rule-based structure is designed for interpretability and can be updated with weights learned via machine learning once labeled handwriting data is available.

3.3.5 Report generation and user feedback

The system utilizes the `xhtml2pdf` library to convert a Jinja2-based HTML template into a downloadable PDF diagnostic report. This report includes the uploaded handwritten image, the OCR result alongside the expected sentence, and a visual comparison highlighting symbol matches and mismatches. It also contains all extracted metrics, such as character error rate (CER), word error rate (WER), and reversal count, along with the final risk classification message. These reports can be archived or shared with clinicians or educators for deeper follow-up. Additionally, this functionality facilitates anonymized data collection to support future supervised learning tasks.

3.4 Behavioral quiz-based screening

To supplement objective modalities, a digital questionnaire targeting behavioral indicators commonly associated with dyslexia was developed. The quiz consists of 9 clinically inspired multiple-choice items (e.g., reversal confusion, difficulty rhyming, inconsistent spelling) derived from established dyslexia

TABLE 1 Risk interpretation based on dyslexia screening score.

Score range	Interpretation
0–6	Unlikely dyslexia - No significant signs.
7–13	Mild signs - Monitor and support.
14–20	Moderate concern - Consider professional screening.
21+	High concern - Seek professional evaluation.

symptom checklists and is given in [Appendix](#). Each question is scored on a 4-point Likert scale (0–3). The total score ranges from 0 to 27, and risk levels are categorized using fixed thresholds ([Table 1](#)).

The quiz is implemented as a React-based web component, allowing for real-time scoring, progress navigation, and retakes. The responses are stored locally for privacy, with potential future integration into the unified backend for multimodal fusion.

Although subjective, the quiz serves as a lightweight behavioral filter to flag potential concerns when objective modules (e.g., eye-tracking) may be unavailable. It is especially useful in remote or early-stage screening, guiding decisions on whether to proceed with deeper testing.

3.5 Support-oriented interventions

While dyslexia is a lifelong neurological condition with no definitive cure, early detection enables timely support to improve reading outcomes and confidence. To bridge the gap between diagnosis and intervention, two supportive modules integrated into this broader framework was proposed:

- (a) **Reading Support System:** To assist children after screening, a lightweight *Reading Support System* was developed that emphasizes syllable-level decoding and pacing adjustments. The system is implemented as a Flask-based web service that interfaces with a React or static front-end, enabling seamless integration into classroom or home environments.

3.5.1 Architecture and workflow

Workflow of the guided syllable-level reading module is given in [Figure 3](#).

3.5.2 Key features

The proposed module incorporates several important characteristics that ensure both usability and efficiency. [Table 2](#) highlights the key features of the syllable-based reading module.

3.5.3 Integration potential

Although currently a standalone tool, the Reading Support System is designed to interoperate with the other modules in this framework. For example, gaze coordinates from real-time eye tracking ([Section 3.1](#)) can trigger automatic syllable advancement, while back-end logs of syllable dwell times could feed into future machine learning models to personalize pacing.

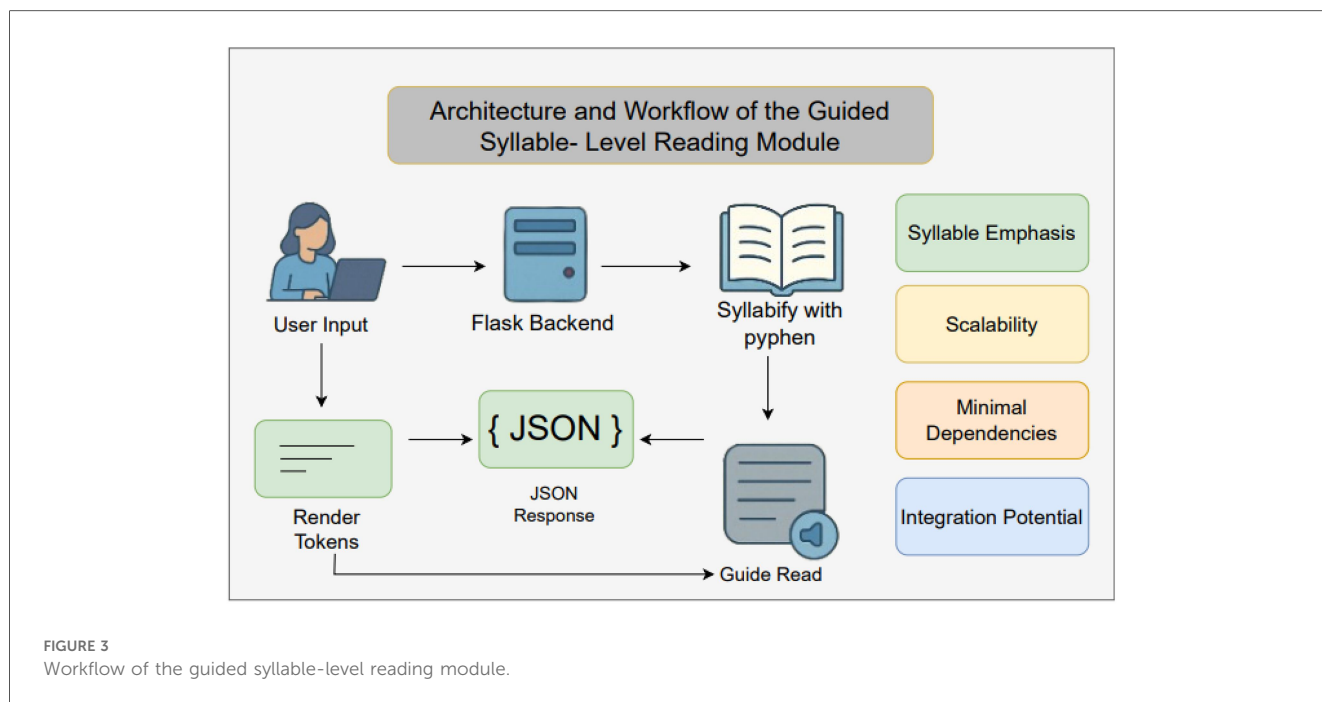


FIGURE 3 Workflow of the guided syllable-level reading module.

TABLE 2 Key features of the syllable-based reading module.

Feature	Description
Syllable emphasis	Reinforces phonological decoding by isolating syllables, reducing cognitive load for struggling readers.
Scalability	Handles passages up to 5,000 words, supporting full-length texts without performance issues.
Minimal dependencies	Requires only Flask, flask_cors, and pyphen, enabling easy deployment even on low-resource systems.

3.5.4 User experience

Clinicians or educators can upload or paste any instructional text; the system instantaneously returns a syllabified version. Children follow on-screen highlights or listen to synchronized text-to-speech (TTS) in future iterations, promoting fluency and confidence.

This intervention bridges the gap between early detection and targeted support, providing an immediately deployable tool to reinforce phonological skills in real time.

- (b) **Gamified Learning Interface:** A game-based environment is being developed in which children engage in letter-sound matching, sequencing tasks, and rhyming puzzles. The gamified design promotes motivation, reduces stigma, and reinforces skills in an adaptive, low-pressure setting.

These modules are not diagnostic but therapeutic, intended to complement the screening tools. Future iterations will integrate user performance into the unified data stream, enabling personalized interventions and progress monitoring over time.

4 Experimental validation and results

The initial testing has demonstrated promising result with controlled dataset as outcomes. The system was able to flag

potential dyslexic traits in children, through the interactive assessment. The eye tracker, Speech to Text, OCR and AI-based Handwriting Analysis are the assessments used in the application. The user feedback indicate that the interface was intuitive and helpful particularly for non-specialist users such as parents or primary school teachers. The application also meets the requirement in the assisting of the dyslexic children. While the system is not a diagnostic tool, it effectively supports early screening objectives.

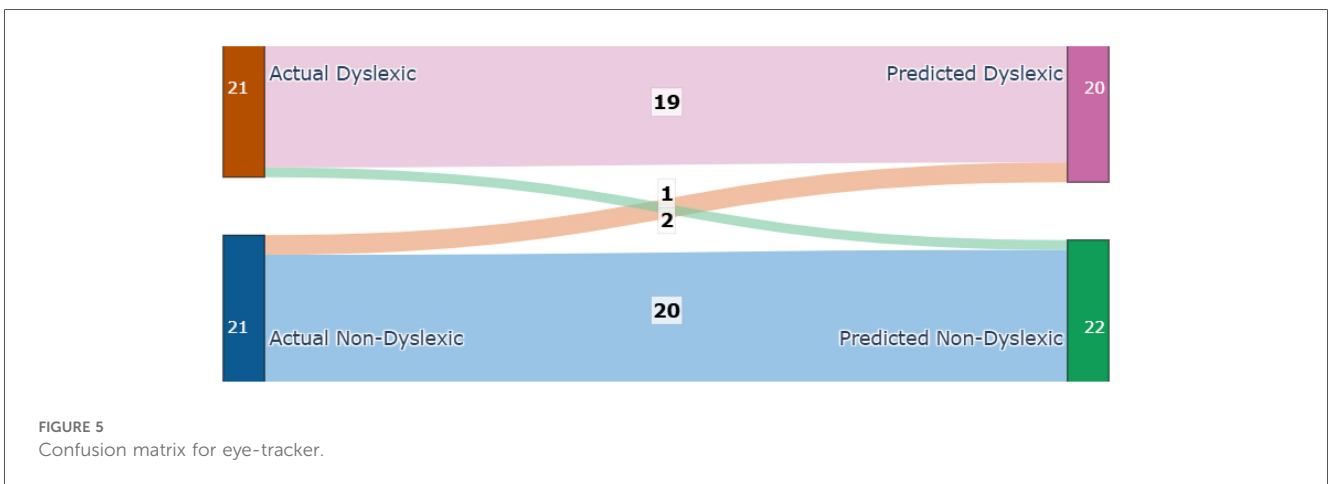
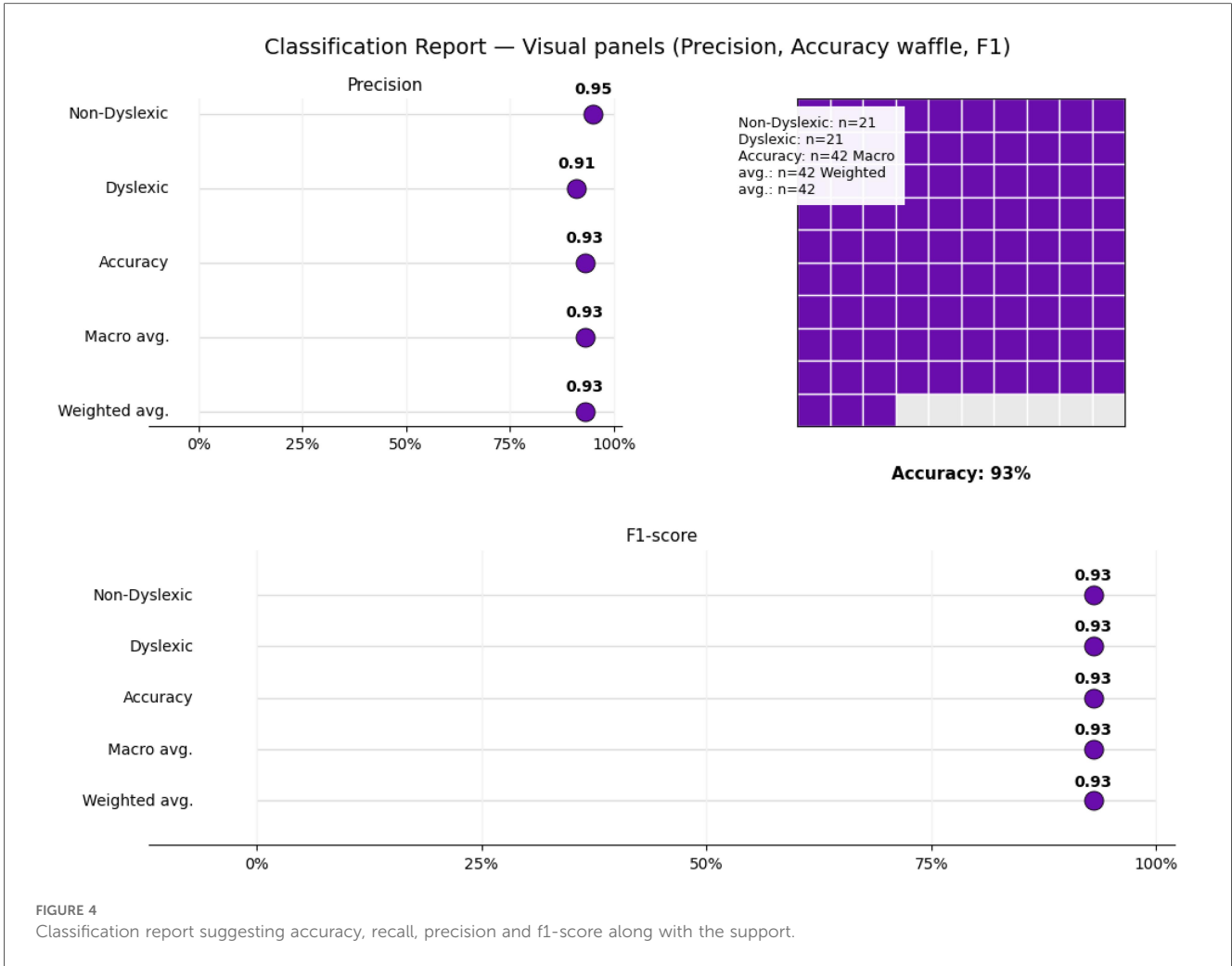
The Akshar Mitra home page presents all six modules-Eye Tracking, Speech Fluency, Handwriting OCR, Behavioral Quiz, Reading Companion and Gamified tools-as interactive cards on a single dashboard. Users can simply click a module’s card to navigate directly to its testing interface and begin assessment.

4.1 Performance analysis

Each module was evaluated on its respective test set using accuracy, precision, recall Figure 4, F₁-score, confusion matrices Figure 5, and ROC analysis Figure 6. The eye-tracking classifier achieved an accuracy of 92.8%, and an F₁-score of 0.93, its ROC curve yielded an AUC of 0.99, with the confusion matrix showing low false positive and negative rates. These results demonstrate that the eye-tracking classifier reliably distinguishes dyslexic from typical readers, with confusion matrices indicating a well-balanced distribution of classification errors across both classes.

4.1.1 Eyes and speech analysis feature

This work describes a new multimodal system developed as a tool for dyslexia screening, which integrates two diagnostic modules-Eye-Gaze Analysis and Speech Analysis-under a unified



digital framework Figure 7, 8. The user protocol is comprised of two sequential tasks: an eye-tracking session, lasting 10 s, that gathers and validates the required number of eye-landmark samples, followed by a speech task lasting 10 s in which the user reads a prescribed sentence. The two modalities are processed separate then the aggregated eye-gaze features are entered into a

trained classifier that outputs an initial estimate, whilst the recorded audio is processed as input to estimate the Word Error Rate (WER) and provide an Automatic Speech Recognition (ASR) transcript. Here, the innovation in the system is the sophisticated combination of these outputs into a “Combined Score” that is rendered along with the individual outputs. By comparing

quantitative indicators (gaze patterns, WER) with qualitative data (playback of the recorded audio, transcripts), this unified system provides a comprehensive, multifaceted interpretation of reading proficiency, thereby providing a more reliable and subtle screening output that could be achieved through unimodal analysis alone.

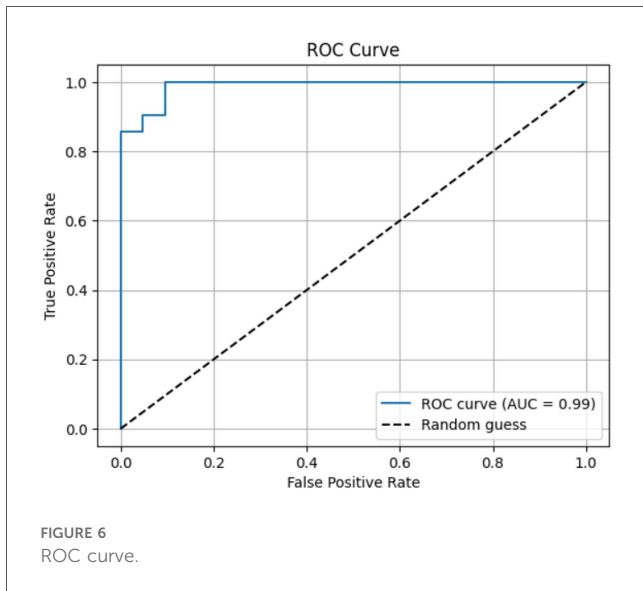


FIGURE 6
ROC curve.

4.1.2 OCR-based dyslexia screening

A specialized software tool analyzes handwriting samples-whose expected text is known-using the EasyOCR library to extract the transcribed string from the input image. The resulting text is aligned with the ground-truth sentence via a sequence-matching algorithm to compute a suite of dyslexia-relevant metrics: Character Error Rate (CER), Word Error Rate (WER), and counts of substitutions, insertions, and deletions (Figure 9). Additionally, the system identifies and tallies letter-reversal errors (e.g., “b” vs. “d”), which are distinctive markers of dyslexic writing. Threshold-based decision logic then classifies samples as “likely dyslexic” when one or more metrics exceed empirically determined cutoffs, flagging the writing for formal follow-up evaluation.

4.1.3 Behavioral quiz outcomes

After the participant completes all questionnaire items (Figure 10), the total Likert-scale score is computed and mapped to one of four risk categories: *Unlikely Dyslexia* (0–6), *Mild Signs* (7–13), *Moderate Concern* (14–20), or *High Concern* (21–27). The interface then displays the numeric score alongside its categorical label, enabling immediate, interpretable feedback.

4.1.4 Reading support tool

A novel software application tailored for dyslexic readers synergistically combines three evidence-based support strategies. First, an integrated Text-to-Speech (TTS) engine vocalizes each

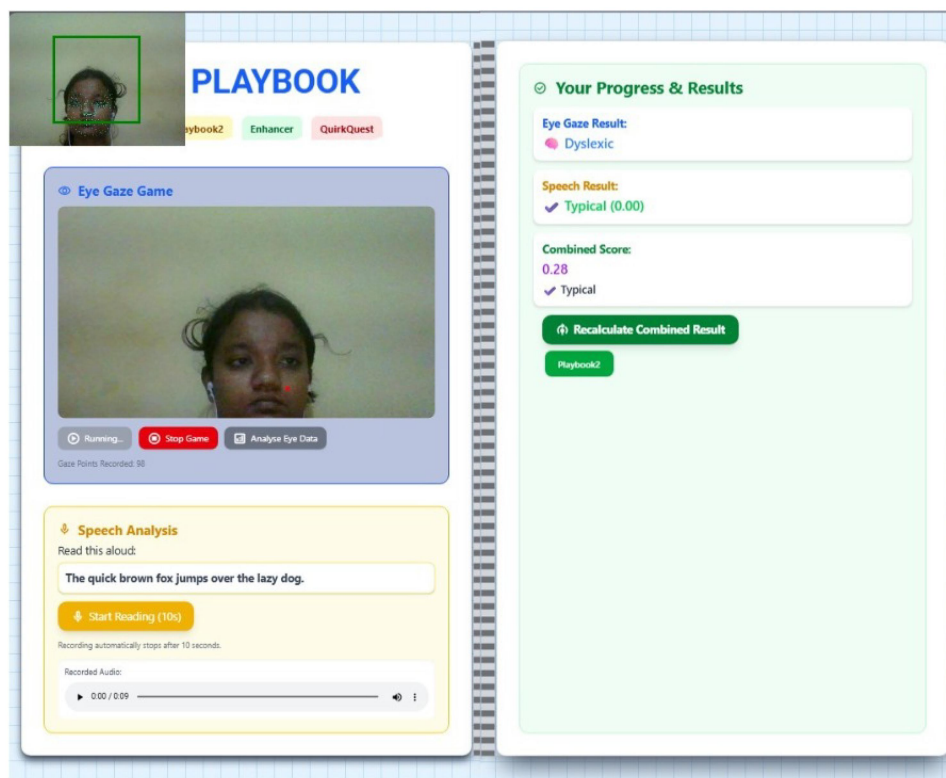


FIGURE 7
Eyes and speech analysis for typical and dyslexic cases.

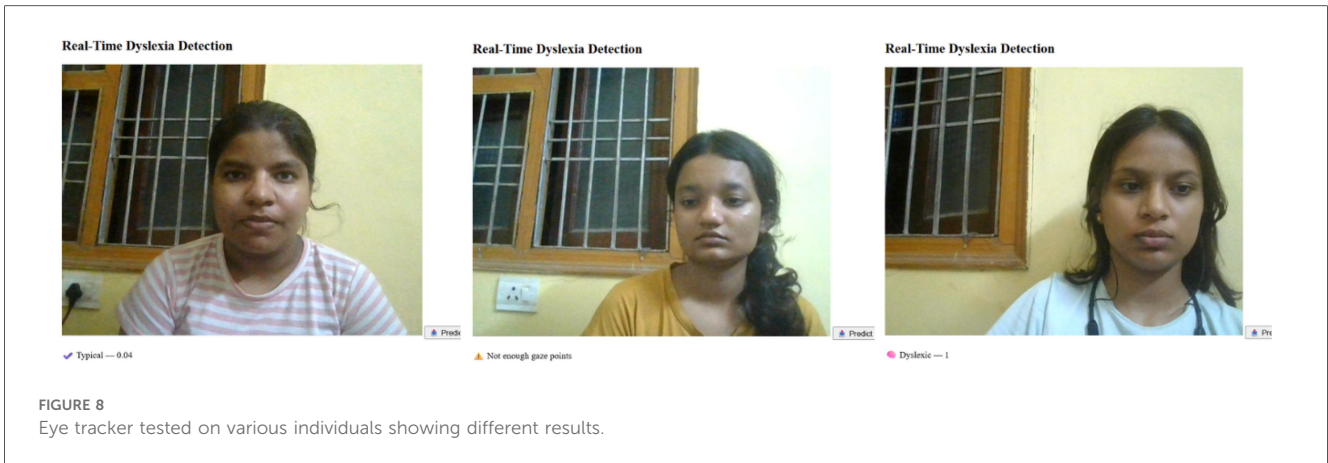


FIGURE 8 Eye tracker tested on various individuals showing different results.

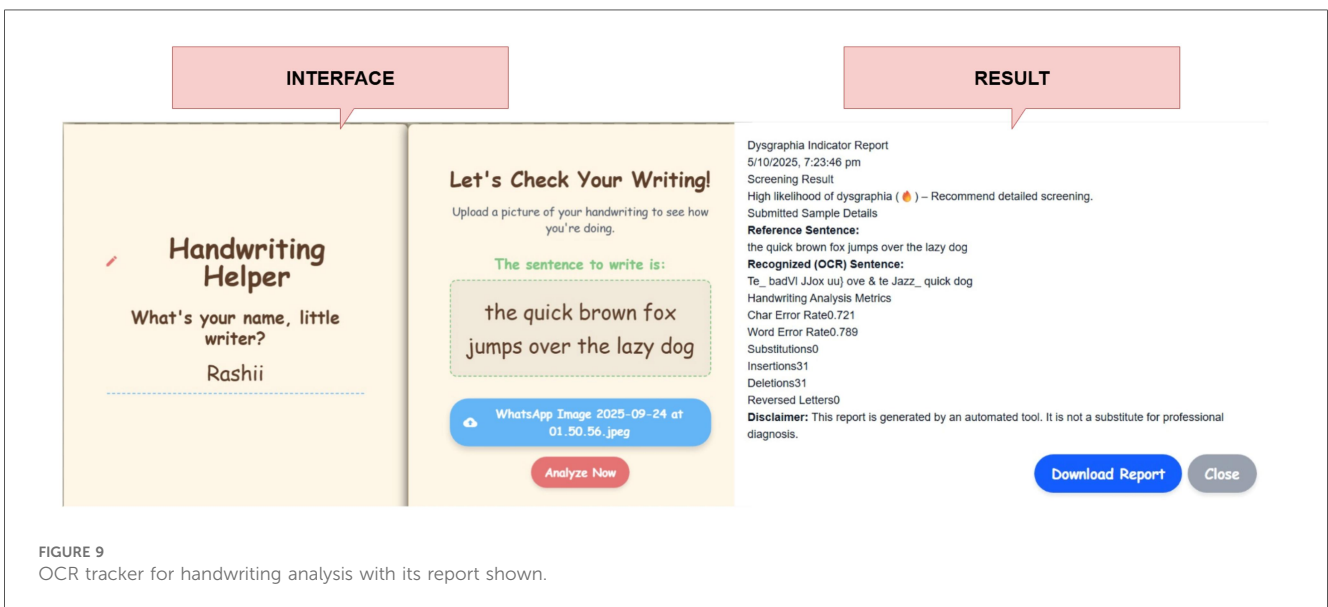


FIGURE 9 OCR tracker for handwriting analysis with its report shown.

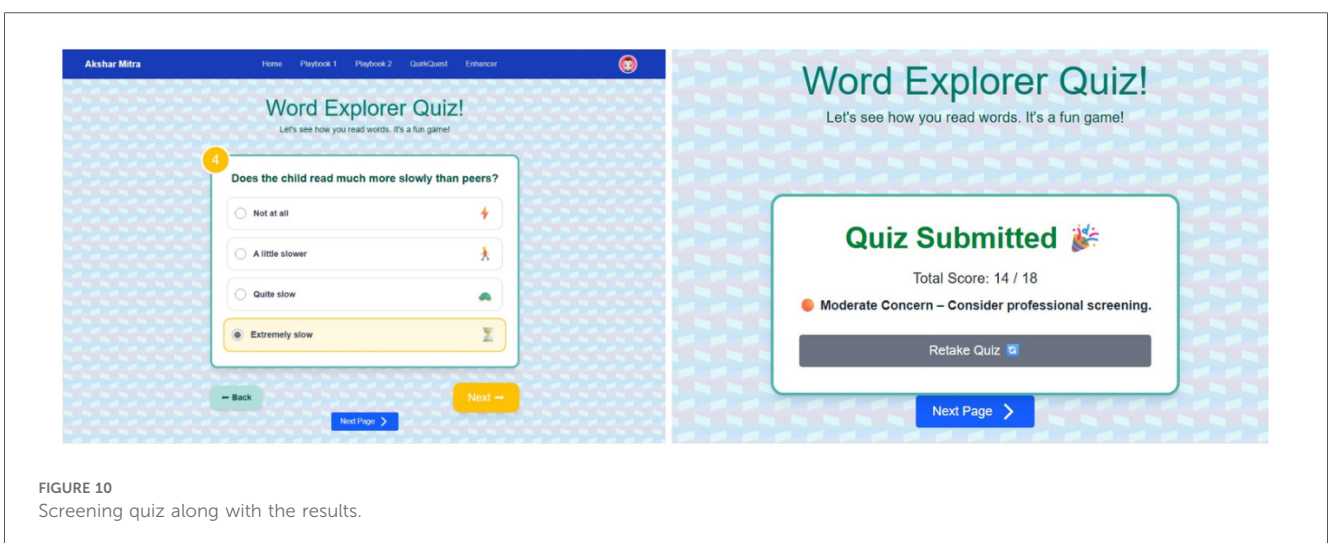


FIGURE 10 Screening quiz along with the results.

word, reinforcing grapheme–phoneme correspondence and aiding auditory processing. Second, a line-by-line reading mode dynamically highlights one text line at a time, minimizing visual tracking errors and enhancing reading flow. Third, an auto-syllabification module segments longer words into

pronounceable syllable units, reducing decoding effort and cognitive load **Figure 11**. By simultaneously targeting auditory, visual, and decoding deficits, the system provides a comprehensive assistive tool that empowers dyslexic individuals to improve reading accuracy and confidence.

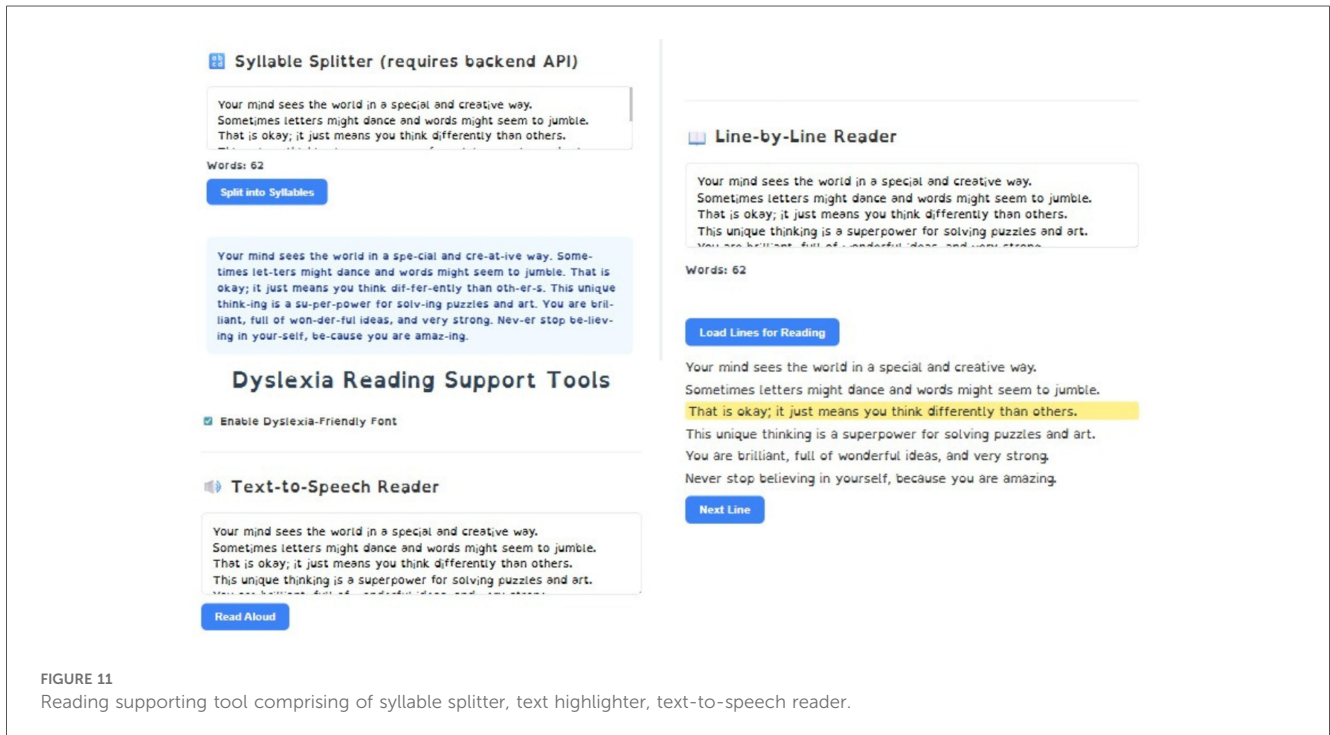


FIGURE 11 Reading supporting tool comprising of syllable splitter, text highlighter, text-to-speech reader.

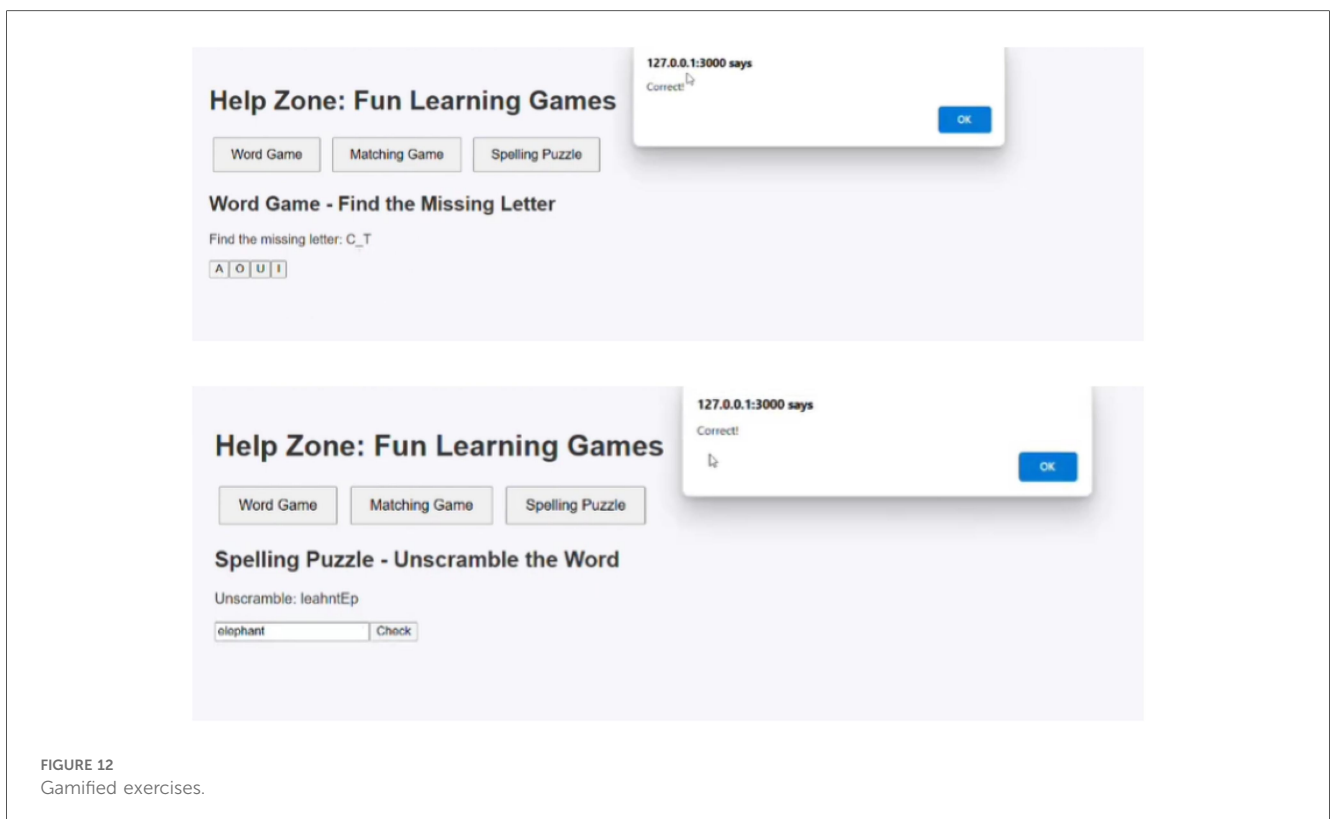


FIGURE 12 Gamified exercises.

4.1.5 Gamified learning outcomes

The platform includes interactive games that promote letter-sound recognition and word formation skills. In the “Missing Letter” activity, children learn to identify letters within words using contextual image clues, enhancing phonemic awareness. The “Word Unscramble” game reinforces spelling and vocabulary by requiring users to reconstruct jumbled words [Figure 12](#). These modules are both educational and diagnostic, capturing subtle learning patterns that support dyslexia assessment.

5 Conclusion

This study introduces a multimodal, technology-assisted framework for the early screening of dyslexia in children, combining eye-movement analysis, handwriting pattern recognition, and audio-based linguistic assessment within an adaptive, interactive environment. The proposed system employs real-time data acquisition and intelligent decision modules to identify early indicators of reading and cognitive difficulties. By unifying behavioral and linguistic modalities, the framework demonstrates the feasibility of data-driven, user-centered screening that can complement conventional diagnostic approaches and empower educators and caregivers with actionable insights.

Nevertheless, the study remains preliminary, as validation is currently constrained by small-scale datasets and limited participant diversity. Future work will focus on large-scale empirical studies (24), cross-linguistic adaptation, and the integration of explainable AI techniques to enhance interpretability and clinical trust. Strengthening ethical safeguards and data transparency will also be central to ensuring responsible deployment. Despite these constraints, this research establishes a meaningful foundation for developing intelligent, inclusive, and ethically aligned assistive technologies that bridge the gap between artificial intelligence and human-centered educational support for children at risk of dyslexia.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Ethics statement

Ethical approval was not required for the study involving humans in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was not required from the participants or the participants’ legal guardians/next of kin in accordance with the national legislation and the institutional requirements. Written informed

consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

VT: Supervision, Conceptualization, Project administration, Writing – review & editing, Methodology, Writing – original draft, Investigation. OA: Writing – original draft, Visualization, Investigation, Resources, Software, Formal analysis, Validation, Writing – review & editing, Data curation. MS: Formal analysis, Writing – original draft, Data curation, Investigation, Resources. RS: Resources, Writing – original draft, Visualization. RB: Writing – original draft, Visualization, Methodology, Data curation. RG: Formal analysis, Writing – review & editing, Data curation, Project administration, Validation. MA: Writing – review & editing, Supervision, Validation. HG: Writing – review & editing, Conceptualization, Investigation, Methodology, Resources, Project administration.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Shaywitz SE. Dyslexia. *N Engl J Med.* (1998) 338(5):307–12. doi: 10.1056/NEJM199801293380507
- Frith U. Beneath the surface of developmental dyslexia. In: Patterson K, Marshall JC, Coltheart M, editors. *Surface Dyslexia*. London: Routledge (2017). p. 301–30.
- Backes W, Vuurman E, Wennekes R, Spronk P, Wuisman M, van Engelshoven J, et al. Atypical brain activation of reading processes in children with developmental dyslexia. *J Child Neurol.* (2002) 17(12):867–71. doi: 10.1177/088307380201701209
- Gabrieli JDE. Dyslexia: a new synergy between education and cognitive neuroscience. *Science.* (2009) 325(5938):280–3. doi: 10.1126/science.1171999
- Shaywitz SE, Shaywitz BA. Reading disability and the brain. *Educ Leadersh.* (2004) 61(6):7–12.
- Shaywitz BA, Lyon GR, Shaywitz SE. The role of functional magnetic resonance imaging in understanding reading and dyslexia. *Dev Neuropsychol.* (2006) 30(1):613–32. doi: 10.1207/s15326942dn3001_5
- Clark R, Williams C, Gilchrist ID. Oculomotor control in children with special educational needs (SEN): the development and piloting of a novel app-based therapeutic intervention. *Health Technol (Berl).* (2021) 11(4):919–28. doi: 10.1007/s12553-021-00538-x
- Al Dahhan NZ, Kirby JR, Brien DC, Munoz DP. Eye movements and articulations during a letter naming speed task: children with and without dyslexia. *J Learn Disabil.* (2017) 50(3):275–85. doi: 10.1177/0022219415609606
- Sedmidubsky J, Dostalova N, Svaricek R, Culemann W. ETDD70: eye-tracking dataset for classification of dyslexia using AI-based methods. In: *International Conference on Similarity Search and Applications*. Springer (2024). p. 34–48.
- Schaur M, Koutny R. Dyslexia, reading/writing disorders: assistive technology and accessibility: introduction to the special thematic session. In: *International Conference on Computers Helping People with Special Needs*. Springer (2024). p. 269–74.
- Al-Barhamtoshy HM, Motaweh DM. Diagnosis of dyslexia using computation analysis. In: *2017 International Conference on Informatics, Health and Technology (ICIHT)*. IEEE (2017). p. 1–7. doi: 10.1109/ICIHT.2017.8058210
- Jing CT, Chen CJ. A research review: how technology helps to improve the learning process of learners with dyslexia. *J Cogn Sci Hum Dev.* (2017) 2(2):26–43. doi: 10.33736/jcshd.510.2017
- Smyrnakis I, Andreadakis V, Selimis V, Kalaitzakis M, Bachourou T, Kaloutsakis G, et al. RADAR: a novel fast-screening method for reading difficulties with special focus on dyslexia. *PLoS ONE.* (2017) 12(8):e0182597. doi: 10.1371/journal.pone.0182597
- Eden GF, VanMeter JW, Rumsey JM, Maisog JM, Woods RP, Zeffiro TA. Abnormal processing of visual motion in dyslexia revealed by functional brain imaging. *Nature.* (1996) 382(6586):66–9. doi: 10.1038/382066a0
- Toki EI. Using eye-tracking to assess dyslexia: a systematic review of emerging evidence. *Educ Sci.* (2024) 14(11):1256. doi: 10.3390/educsci14111256
- Asvestopoulou T, Manousaki V, Psistakis A, Smyrnakis I, Andreadakis V, Aslanides IM, et al. DysLexML: screening tool for dyslexia using machine learning. *arXiv:1903.06274* (2019). Available online at: <http://arxiv.org/abs/1903.06274> (Accessed August 17, 2025).
- Cogan K, Ngo VM, Roantree M. Developing a dyslexia indicator using eye tracking. In: *International Conference on Artificial Intelligence in Medicine*. Springer (2025). p. 105–9.
- Temple CM, Marshall JC. A case study of developmental phonological dyslexia. *Br J Psychol.* (1983) 74(4):517–33. doi: 10.1111/j.2044-8295.1983.tb01845.x
- Park HJ, Takahashi K, Roberts KD, Delise D. Effects of text-to-speech software use on the reading proficiency of high school struggling readers. *Assist Technol.* (2017) 29(3):146–52. doi: 10.1080/10400435.2016.1161732
- Gemelli A, Marinai S, Vivoli E, Zappaterra T. Deep-learning for dysgraphia detection in children handwritings. In: *Proceedings of the ACM Symposium on Document Engineering 2023*. ACM (2023). p. 1–4. doi: 10.1145/3573128.3609351
- Singh A, Garg SK. Comparative study of optical character recognition using different techniques on scanned handwritten images. In: *Micro-Electronics and Telecommunication Engineering: Proceedings of 6th ICMETE 2022*. Springer (2023). p. 411–20.
- Dostalova N, Svaricek R, Sedmidubsky J, Culemann W, Sasinka C, Zezula P, et al. *ETDD70: Eye-Tracking Dyslexia Dataset*. Zenodo (2024). doi: 10.5281/zenodo.13332134.
- Lyytinen H, Erskine J, Tolvanen A, Torppa M, Poikkeus AM, Lyytinen P. Trajectories of reading development: a follow-up from birth to school age of children with and without risk for dyslexia. *Merrill Palmer Q.* (2006) 52(3):514–46. doi: 10.1353/mpq.2006.0025
- Verpalen A. *Dyslexia in a Multicultural Context* (doctoral thesis). Tilburg University (2017).
- Lodygowska E, Chec M, Samochowicz A. Academic motivation in children with dyslexia. *J Educ Res.* (2017) 110(5):575–80. doi: 10.1080/00220671.2016.1234567
- Nagaonkar S, Sharma A, Choithani A, Trivedi A. Benchmarking vision-language models on optical character recognition in dynamic video environments. *arXiv:2502.06445* (2025). Available online at: <https://arxiv.org/abs/2502.06445> (Accessed August 17, 2025).

Appendix

Quiz questions

The Behavioral Quiz module (Section 3.4) uses the following nine items. Respondents select one option per question on a 4-point scale (0–3).

1. Does your child tend to guess words instead of sounding them out while reading?
 - a. Never (0)
 - b. Every now and then (1)
 - c. Quite often (2)
 - d. Almost always (3)
2. Does your child find it hard to recognize common sight words (like “the,” “and,” “said”)?
 - a. Not at all (0)
 - b. Just a little (1)
 - c. Often (2)
 - d. All the time (3)
3. When reading aloud, does your child pause frequently or skip over words or even whole lines?
 - a. Not really (0)
 - b. Sometimes (1)
 - c. Quite frequently (2)
 - d. Very often (3)
4. Does your child mix up letters like “b” and “d” or “p” and “q”?
 - a. Rarely (0)
 - b. Occasionally (1)
 - c. Frequently (2)
 - d. Almost always (3)
5. Does your child misspell simple words in ways that seem unpredictable or inconsistent?
 - a. No (0)
 - b. Sometimes (1)
 - c. Often (2)
 - d. Very often (3)
6. Does your child find it difficult to rhyme words or pick out individual sounds in words?
 - a. Not at all (0)
 - b. A little bit (1)
 - c. Quite a bit (2)
 - d. A lot (3)
7. Does your child sometimes mix up sounds or syllables when speaking?
 - a. Not really (0)
 - b. Rarely (1)
 - c. Sometimes (2)
 - d. Often (3)
8. Does your child struggle to remember things that come in a sequence, like the alphabet or days of the week?
 - a. No (0)
 - b. Occasionally (1)
 - c. Often (2)
 - d. Very often (3)
9. Does your child have a hard time staying focused when reading or writing?
 - a. No (0)
 - b. Occasionally (1)
 - c. Frequently (2)
 - d. Almost constantly (3)